AGENT AND METHOD FOR MODULATION OF CELL MIGRATION

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of provisional patent applications 60/087,170, filed May 29, 1998, and 60/129,023, filed April 13, 1999, each of which is incorporated herein by reference in its entirety.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

To be determined.

BACKGROUND OF THE INVENTION

Cell migration, particularly migration of cancerous cells and nerve cells, is not well understood, nor are the factors that affect cell migration and tissue shaping *in vivo*. There is a need in the art to identify and exploit such factors, including but not limited to those involved in normal or abnormal organogenesis. The art also lacks efficient systems for evaluating therapeutic modulators of such functions *in vivo* and lacks diagnostic methods for assessing the ability of a cell or cell mass to migrate *in vivo*.

Organogenesis processes in vertebrates proceed in a manner similar to those observed in the common laboratory nematode *C. elegans*. As such, the generation of *C. elegans* gonadal structures can serve as a simple system for investigating developmental morphogenetic processes shared by higher and lower organisms.

In one common morphogenetic process, a tissue bud extends to form an elongate tube with a proximal to distal axis. An emerging theme in bud extension is the presence of specialized regulatory cells at the bud tip that govern elongation. In vertebrate development, this process is

seen in extension of the limb (Johnson and Tabin, 1997; Martin, 1998), ureter (Vainio and Muller, 1997), and lung branches (Hogan, 1998). In the *C. elegans* gonad, long "arms" develop by elongation of buds originating from a gonadal primordium. Each gonadal arm possesses a single "leader cell" that serves this regulatory role (Kimble and White, 1981). The biology of distal tip cell migration during gonadogenesis is known to one skilled in the art of *C. elegans* developmental biology. Indeed, the *C. elegans* gonadal leader cells are among the best defined cells that regulate bud elongation, and therefore serve as a paradigm for investigating this common morphogenetic process.

A second common morphogenetic process of organogenesis is the formation of a complex, differentiated epithelial tube. Formation of a complex epithelial tube can involve an initial condensation of mesenchymal cells, followed by epithelialization, lumen formation, and differentiation into modular units. Vertebrate examples include the kidney tubules (Vainio and Muller, 1997) and heart tube (Fishman and Olson, 1997). Similarly, during *C. elegans* gonadogenesis, cells coalesce to form a compact larval structure called the somatic gonadal primordium (SGP). Following formation of this primordium, cell division and differentiation are accompanied by epithelialization and lumen formation to form a complex tube composed of distinct modular units: the uterus, spermathecae and sheaths in hermaphrodites, and the seminal vesicle and *vas deferens* in males (Kimble and Hirsh, 1979).

Previous studies have identified several genes in *C. elegans* that influence gonadal morphogenesis. One group of such genes includes *unc-5*, *unc-6*, and *unc-40*, which control the direction of leader cell migration (Hedgecock et al, 1990). Normally, leader cells migrate in one direction, then move dorsally, and finally move in the opposite direction to generate a reflexed gonadal arm. In the absence of *unc-5*, *unc-6*, or *unc-40*, the leader cells fail to turn dorsally. Another gene, *ced-5*, causes the

leader cell to makes extra turns or stop prematurely (Wu and Horvitz, 1998). Therefore, in these mutants, the leader cells migrate, but do not navigate correctly, which results in a failure of the gonadal arms to acquire their normal U-shape. In addition to these genes, others are required for specification of cell fates and also influence morphogenesis (*lin-12:* Greenwald et al., 1983, Newman et al., 1995; *lin-17:* Sternberg and Horvitz, 1988; *lag-2:* Lambie and Kimble, 1991; *ceh-18:* Greenstein et al., 1994, Rose et al., 1997; *lin-26:* den Boer et al., 1998).

A known *C. elegans* genetic locus, *gon-1*, defined by one or more mutants, is essential for extension of gonadal germline arms, but is not responsible for signaling the germline to proliferate. In *C. elegans* hermaphrodites, GON-1 is required for migration of two distal tip cells to produce two elongated tubes, whereas in males, *gon-1* activity is required for migration of a single linker cell to produce a single elongated tube. In *gon-1* mutant hermaphrodites, the leader cells are born normally in the somatic gonadal cell lineage and function normally to promote germline proliferation, but they fail to migrate and do not support arm extension. Similarly in males, the leader cell does not move and no arm extension occurs. The *gon-1* locus has not heretofore been mapped with particularity to a nucleic acid coding sequence.

Clarification of the genetic basis for *C. elegans gon-1* activity would permit one to apply molecular tools to the study of cell migration in a convenient system. It would be particularly advantageous to find that the *gon-1* locus encodes a protein having structural relationship to proteins of species that are not readily studied in the laboratory, since one would be able to evaluate those proteins in the convenient *C. elegans* system. Such a system would also provide a means for evaluating agents that can modulate the activity of such genes and proteins and would both facilitate understanding the factors involved in cell migration.

BRIEF SUMMARY OF THE INVENTION

In one aspect, the invention can be an isolated polynucleotide coding sequence that encodes a protein the includes both a metalloprotease domain and at least one thrombospodin type 1 domain, where the protein can direct either cell migration or tissue shaping in an analytical system in a target organism as disclosed herein. In another aspect, the invention can also be a variant of the isolated polynucleotide coding sequence that encodes a protein that shares at least 20%, more preferably 50%, still more preferably 70% and most preferably 80% amino acid sequence identity (using GCG Pileup program) with any of the foregoing in the metalloprotease and thrombospondin type 1 domains while also comprising the amino acids of those domains known to those skilled in the art to be required for protein activity. A suitable variant polynucleotide can hybridize under stringent hybridization conditions known to those skilled in the art to a polynucleotide sequence that encodes a protein that can direct cell migration or tissue shaping in the target organism. In one embodiment, a variant polynucleotide can hybridize under stringent hybridization conditions to a *C. elegans gon-1* coding sequence. The variant polynucleotide sequence can be a polynucleotide obtained from an organism or can be a mutated version of any polynucleotide sequence noted above. The variant polynucleotide can encode a protein that is identical or altered relative to the wild-type *C. elegans* GON-1 protein. The encoded protein can have enhanced or reduced activity *in vivo* relative to GON-1.

In a related aspect, a polynucleotide coding sequence that encodes a protein having structural and functional similarity with a wild-type or altered migration or shaping protein can also be substituted, in whole or in part, with structurally related or unrelated sequences to encode a heterologous protein or a chimeric protein in the disclosed system, as detailed below.

Applicants herein disclose that the *Caenorhabditis elegans gon-1* activity is encoded by a polynucleotide coding sequence (*gon-1*; SEQ ID NO:1) that encodes an essential protein (GON-1; SEQ ID NO:2) that directs migration of a growing gonadal tube through surrounding basement membranes during gonadogenesis in the nematode and also controls gonadal shape and organ localization.

The migration directing ability and tissue shaping ability are separable and depend upon whether the *gon-1* coding sequence is expressed in distal tip cells or in muscle cells, respectively. In wild-type *C. elegans*, a gonad of normal shape is produced when *gon-1* is expressed in both cell types. Accordingly, one aspect of the invention can also a method for shaping a tissue by selectively expressing a protein associated with both tissue elongation and tissue expansion. GON-1 shares significant amino acid identity with proteins that have been noted in other species.

In a related aspect, the invention can be an isolated and substantially purified preparation of a GON-1 protein, an altered GON-1 protein, a heterologous protein, a chimeric protein, or a variant thereof (referred to herein as "an MPT protein", for reasons discussed below), which can be a target for *in vivo* screening of putative therapeutic modulators, or can be assayed in a diagnostic method for assessing the ability of a cell or cell mass to migrate *in vivo*, or can be exploited as a therapeutic agent to modulate (increase or decrease) *in vivo* cell migration.

One skilled in the art will appreciate that the nucleotide coding sequences and encoded amino acid sequences that fall within the scope of the invention are also subject to natural variation or intentional manipulation (e.g., changes in the nucleotide or amino acid sequence) in ways that do not affect the ability to function as described herein. One skilled in that art also understands that the applicants cannot provide a complete list of nucleotide coding sequences and amino acid

sequences that can function in the methods of the invention. However, in view of the high level of understanding in the art about the amino acids required for activity of proteins that comprise a metalloprotease domain and proteins that comprise a thrombospondin domain, applicants maintain that a skilled artisan can readily determine whether a protein contains both domains. Stöcker, W. et al., "The metzincins - Topological and sequential relations between the atacins, adamalysins, serralysins, and matrixings (collagenases) define a superfamily of zinc-peptidases," Protein Science 4:823-840 (1995), Rawlings, N.D. and A.J. Barrett, "Evolutionary families of metallopeptidases, Methods in Enzymology 248:183-228 (1995), and Adams, J.C. et al., The Thrombospondin Gene Family, R.G. Landes Company, Austin, TX (1995), all incorporated herein by reference in their entirety, provide sufficient guidance to permit those in the art to establish whether a protein comprises both a metalloprotease and a thrombospondin domain.

The invention is further summarized in that an antibody can be produced against characteristic epitopes of any of the foregoing proteins using standard methods. The antibody can be used both diagnostically to ascertain the presence of an MPT protein, or therapeutically to interfere with activity of the MPT protein.

The present invention is also summarized in that an animal that contains a *gon-1* allele (or homolog or variant thereof) is a convenient screening tool for finding modulators of cell migration. The present invention is thus further summarized in that a method for identifying modulators of the disclosed MPT proteins includes the steps of treating a target organism having a cell that can migrate or be shaped when under control of an MPT protein with at least one potential modulator of migration or shaping and observing in the treated target organism a change in migration or shaping of the cell or tissue attributable to the presence of a modulator. In a

preferred embodiment, the cell is a developing gonadal cell in *C. elegans*, although other cells or organs may be similarly regulated by MPT proteins in other organisms.

The ability of the MPT protein to direct a cell or tissue under its influence to migrate or be shaped can be modulated (increased or decreased) in a variety of ways, such as by altering the migration protein's primary, secondary, or tertiary structure, by altering the location or amount of the protein in an organism, by altering the transcriptional or translational regulation of the gene that encodes the protein, or by providing the organism with an agonist or antagonist molecule in an amount sufficient to interact with the MPT protein so as to increase or decrease the ability of the protein to direct migration or shaping.

In a related method, one can also identify nucleic acid sequences required or desired for migration or shaping of such a cell, by treating a target organism with an agent that affects the polynucleotide sequences of the target organism that encode the MPT protein or that participate in regulating expression of the MPT protein, and then identifying sequences affected by the treatment. The sequences identified in the method can be either complete or partial coding sequences or can be regulatory sequences.

It is an object of the present invention to identify a protein and nucleotide sequence encoding same that directs migration or shaping of a cell or tissue.

It is another object of the present invention to provide a method for modulating cell migration or shaping.

It is yet another object of the present invention to provide a system and method for screening putative modulators of migration or shaping of cells or tissues.

It is an advantage of the present invention that agents having a putative effect upon migration or shaping can be screened in a convenient model system rather than in a vertebrate organism.

Other objects, features and advantages of present

invention will become apparent upon consideration of the following detailed description taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

Fig. 1A depicts a schematic map of the *gon-1* locus in *C. elegans* from which the gene was cloned and shows the exon-intron structure of *gon-1*.

Fig. 1B shows a schematic map of *C. elegans* GON-1, the location of five protein-truncating stop mutants in GON-1 and a comparison to the protein structures of the murine ADAMTS-1 protein, and the bovine procollagen-I N-proteinase (PN1P) protein. From left to right, GON-1 includes a prodomain, a metalloprotease domain, a first cysteine rich region, a thrombospondin type I motif, a second cysteine rich region, and a plurality of thrombospondin type I-like motifs. The five mutants are identified as *q518 (aa591 TGG->TGA)*, *e2551 (aa1069 TGG->TAG)*, *e2547 (aa1229 TGG->TGA)*, *q18 (aa1234 TGG->TAG)* W->stop, and e1254 (aa1345 CGA->TGA) R->stop).

Fig. 1C compares the *C. elegans* GON-1 amino acid sequence to sequences of the ADAMTS-1 and PN1P proteins. In the metalloprotease domain, amino acids important for enzymatic activity are marked by an asterisk (*). Three conserved histidines (GON-1, aa 424, 428, 434) bind a catalytically essential Zn^{+2} ion in well characterized metalloproteases, while a glutamic acid residue (GON-1, aa 425) is thought to be directly involved in cleavage (Stöcker et al, 1995). In addition, two conserved glycines and a downstream methionine seem to be important for structure of the active site. GON-1 bears one of the glycines (aa 427) and the methionine (aa 454), but the second glycine is changed to serine in GON-1 (aa431). In the canonical TSPt1 domain, amino acids conserved in vertebrate TSP type-1rpeats are shown by a plus (+). The mutation, *gon-1(q518)*, is marked by an inverted triangle

(V). For the TSPt1-like repeats, only 2 of the 17 are shown. The consensus sequence for these repeats is: $W\text{-}X_{4\text{-}5}\text{-}W\text{-}X_2\text{-}\ CS\text{-}X_2\text{-}CG\text{-}X_{4\text{-}5}\text{-}X\text{-}G\text{-}X_3\text{-}R\text{-}X_3\text{-}C\text{-}X_{4\text{-}27}C\text{-}X_{8\text{-}12}\text{-}C\text{-}X_{3\text{-}4}\text{-}C$. Because only the first two TSPt1-like motifs are shown, the other mutations are not indicated in this figure.

Fig. 2A depicts normal morphogenesis of the *C. elegans* hermaphrodite gonad.

Fig. 2B shows that arm extension does not occur in *gon-1* mutants and that the gonad develops as a disorganized mass of somatic and germline tissues. Similarly, in males, the *gon-1* mutant gonad is severely disorganized and does not acquire its normal shape.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The existence of a protein in *C. elegans* required for cell migration or shaping has not heretofore been known, nor has any function been previously ascribed to a protein encoded by the designated sequence. The inventors have determined that a functional GON-1 protein is required for migration of the regulatory cells that lead the developing gonad organ during its migration. GON-1 is also involved in shaping tissues such as gonads. By appreciating the role of GON-1 (and the *gon-1* gene) and its relationship to a related gene that is upregulated in a metastatic tumor cell, the inventors have identified a gene and protein believed to be fundamental in the process of normal and abnormal cell migration and tissue shaping. The gene and protein, and related genes and proteins, can be utilized in the methods of the invention as described herein. References herein to influencing cell migration are also intended to encompass shaping of tissues or organs. Likewise, references to a migration protein encompass proteins of the same class that can also be used in methods for shaping tissues or organs.

Generally speaking, the methods of the present invention permit one to identify agents that modulate cell migration or tissue shaping *in vivo* or *in vitro*. One can

treat target organisms with panels of polynucleotides, proteins, sugars, lipids, organic molecules, other chemicals, synthetic or natural pharmaceutical agents or other agents to determine whether any agent affects activity of an MST protein. This list is necessarily incomplete, since one cannot predict in advance which agents will be effective. However, applicants have enabled a system for screening panels of putative agents, in accord with the common practices of pharmaceutical companies that typically screen thousands of compounds against a test system in an effort to reveal preferred agents. Candidate agents likely to modulate MPT proteins in the disclosed system include tissue inhibitors of metalloproteases and pharmaceutical metalloprotease inhibitors or enhancers such as those from British Biotech. Inhibitors or enhancers of thrombospondin activity are also good candidate agents.

Agents so identified can be used therapeutically to enhance or inhibit cell migration or to influence tissue shape. Agents having an adverse or inhibiting or knock-out effect upon activity of a migration protein can also be used in a method for biocontrol of animals that employ the migration protein in gonadal development, where the method includes the step of exposing a developing animal to an amount of the agent effective to prevent gonadal development such that the animals are rendered sterile. While this biocontrol method is particularly envisioned for use in nematodes, it may be applicable to other animals as well, since genes related structurally and functionally to *gon-1* are known to exist in animals as diverse as nematodes, cattle and humans.

Using the invention one can also identify polynucleotide sequences including coding and regulatory sequences that affect activity of a migration protein. For example, null or so-called reduced activity mutants can be mutagenized and assayed for activity-restoring, activity-inhibiting or activity-enhancing changes. By extension, one can perform comparable screens *ad infinitum* on

sequences identified in this manner, to obtain still more sequences that have an indirect effect on migration activity. After identifying such sequences in a target organism, one can obtain homologous polynucleotides from other organisms by screening nucleic acid libraries under stringent hybridization conditions in a manner known to those skilled in the art.

A method for evaluating putative modulators of cell migration preferably employs a nematode as a target organism. The methods may be advantageously practiced using a nematode that comprises a migration protein as described herein, or a mutant nematode that either lacks a migration protein or contains a migration protein having reduced activity. The protein can be encoded by wild-type *C. elegans gon-1* (disclosed herein), by a mutant that confers upon the nematode an enhanced or reduced sensitivity to modulators, by a transgene from another organism, in whole or in part, or by a variant of any of the foregoing. Nematodes are desirable target organisms, in general, because they are easy to grow and maintain, and easy to assay, particularly because they are transparent.

Nematodes are also particularly desired because the powerful techniques of reverse genetics can be employed. One can also target specific *C. elegans* sequences for mutation or RNA-mediated interference (a technique used to transiently knock genes out by RNA injection) to identify nucleic acid and protein sequences that have a direct inhibitory or enhancing effect on *gon-1* activity.

With the identification of the *gon-1* gene and GON-1 protein in *C. elegans* and the discovery of homologous genes in other species, the functions of migration proteins can be analyzed *in vivo* during organogenesis using the full force of molecular genetics available in that system. Such functions can include, but may not be limited to cell migration, basement membrane remodeling, and tubular organ formation.

Although the system is exemplified in *C. elegans*, a

free-living (i.e., non-parasitic) nematode, those skilled in the art can develop similar systems operating on the same principles without undue experimentation in other convenient organisms, including other nematodes including, without limitation, *C. briggsae*, or in, for example, *Drosophila*, or other organisms conveniently studied in the laboratory. To do so, one would only need to identify the homolog of *gon-1* in such an organism, using standard molecular biological methods and then screen for related genes, proteins and other factors as described herein. One could also use such systems in other animals to study transgenes in ways comparable to those described herein. Those skilled in the art can produce transgenic animals of many species without undue experimentation.

In the method, a putative modulator is provided to the target organism, for example, by adding it to the growth media, by injecting it into the organism or by gene transformation technology. The effects of said modulator can be assessed either by screening for changes in cell migration or by genetic selection for fertile animals. The assessment methods are known to those skilled in the art. *Caenorhabditis elegans*: Modern Biological Analysis of an Organism, Methods in Cell Biology, volume 48, Epstein, H. F. and D. C. Shakes, eds., Academic Press (1995), incorporated herein by reference in its entirety, describes suitable methods and conditions for growing and monitoring *C. elegans*.

C. elegans GON-1 is characterized by a multi-domain structure that includes several known motifs. GON-1 protein is a secreted metalloproteinase that lacks a transmembrane domain and possesses a predicted metalloprotease domain between amino acids 269-456. The metalloprotease enzymatic activity is essential for GON-1 function; proteins that might be cleaved by this metalloproteinase include components of the basement membrane and other proteins that modulate migration. The metalloprotease domain shares sequence similarity with other metalloproteinase enzymes.

In addition to its metalloprotease domain, GON-1 possesses a series of consecutive motifs that are related to, but variants of, the thrombospondin type 1 (TSPt1) repeats (Fig. 1B,C). The most N-terminal TSPt1 repeat bears the hallmarks of this type of motif in vertebrate thrombospondins (15/16 of the consensus amino acids, + in Fig. 1C) (Adams et al., 1995), whereas the remaining 17 repeats are less similar and define a TSPt1-like variant. Proteins that might interact with this domain include proteins that modulate migration, including but not limited to components of the basement membrane.

GON-1 is similar to members of the reprolysin subfamily (Rawlings, N.D. and A.J. Barrett, "Evolutionary families of metallopeptidases, Methods in Enzymology 248:183-228 (1995), incorporated herein by reference in its entirety). At the N-terminal border of the metalloprotease domain, there is a potential furin cleavage site (Fig. 1C) (Pei and Weiss, 1995; Pei and Weiss, 1996). GON-1 and the reprolysins share a common zinc binding active site with the larger metzincin superfamily (Stöcker et al., 1995). Amino acid conservation within the active site together with the known crystal structure of several superfamily members reveals those amino acids essential for enzymatic activity (marked by asterisks in Fig. 1c) (ibid). GON-1 has all amino acids implicated in catalysis and all but one implicated in structure of the active site.

Wild-type *C. elegans* GON-1 (SEQ ID NO:2) is suitable for use in the methods of the present invention, although a skilled artisan can replace the *C. elegans gon-1* coding sequence with a sequence that encodes all or part of a homologous protein, using the standard tools available to a molecular biologist. This mixing and matching can increase or decrease the activity of the encoded chimeric protein. As described elsewhere herein, it can be desirable to provide a system having reduced or enhanced migration activity, or even no migration activity, depending upon whether one is evaluating agents that enhance or inhibit

migration. Increased gene activity is characterized either by increased gonadal arm extension, increased compactness of gonadal tissue, or fertility. Decreased gene activity is assayed either by decreased gonadal arm extension, decreased compactness of gonadal tissue or sterility. Certain specific activity-reducing mutations in *gon-1* are described in the Examples.

Sequences with related structures have already been isolated from vertebrate organisms, but no related invertebrate sequence is known to the inventors. Still other related metalloprotease proteins (and polynucleotide sequences encoding same) will be isolated from vertebrate and invertebrate organisms. While the *C. elegans gon-1* protein includes 17 thrombospondin domains, the bovine and murine homologs include only 2 such domains. Other known members of the family also have one canonical TSPt1 repeat, can contain at least one TSPt1-like variant repeat, and contain two conserved cysteine rich regions. Based on this conserved architecture, we suggest the name MPT (for MetalloProtease with TSP1 repeats) for the family.

While the *in vivo* functions of these proteins may differ from that of *C. elegans* GON-1, these proteins are expected to function in place of GON-1 in whole or in part in the disclosed methods. All such homologs from other vertebrate and invertebrate organisms (and the polynucleotide sequences that encode such homologs), variants thereof, and chimerics that incorporate portions thereof, whether obtained naturally or induced in the laboratory using the tools available to a molecular biologist, are considered to be useful in the present invention. In particular, functional domains, such as the metalloprotease domain, can be swapped into corresponding domains in *gon-1*.

The amino acid sequences of GON-1, ADAMTS-1 and bovine PN1P are compared in Fig. 1C. The additional thrombospondin domains of GON-1 not found in ADAMTS-1 or PN1P are not shown in Fig. 1C. Those portions of GON-1

that have no obvious relationship to known motifs are conserved among the family of GON-1 homologs. The GON-1 protein shows significant sequence similarity to the bovine procollagen-1 N-proteinase (P1NP), to the murine ADAMTS-1 protein, and to a pair of human aggrecan-degrading metalloprotease-encoding sequences described in International Patent Application Number PCT/US98/15438, published on February 4, 1999 as International Publication No. WO 99/05291, incorporated herein by reference in its entirety. Another human homolog which has significant identity to the bovine P1NP has Genbank accession number d1021662.

Bovine P1NP can proteolyze the N-terminal propeptide from collagen I (Colige et al., 1995, Colige et al., 1997). Metalloprotease activity is required for GON-1 function and suggest that, like P1NP, it may cleave components of the extracellular matrix. Murine adamts-1 expression correlates with tumor cell progression (Kuno et al., 1997). The murine ADAMTS-1 protein is found in an advanced cachexogenic murine tumor cell. Human aggrecanase has been associated with arthritis in humans. Given the role of GON-1 in regulating cell migration of the C. elegans leader cell, we suggest that MPT proteins may be involved more generally in cell migrations that must pass through extracellular matrix and that, in cancerous tissues, loss of MPT regulation may promote metastasis. The percent identity of the identified domains of *C. elegans* GON-1 with the bovine and murine proteins is shown in Fig. 1B.

Changes can be made in any of the foregoing at the nucleic acid level in a manner known to those skilled in the art, by, for example, removing a section of the coding sequence, interrupting the coding sequence with an additional sequence, rearranging at least one section of the gene, or by providing in the sequence other changes that can include but are not limited to point mutations that either truncate the protein or disable an active site in the protein encoded by the altered polynucleotide.

Changes can also be made by altering the transcription or translation of the gene that encodes the migration protein by altering in a manner known to the art the upstream and/or downstream regulatory sequences that the surround the gene. Likewise the translation-regulating elements of an mRNA encoding the migration protein can also be altered to affect the stability or location of the mRNA. An antisense RNA can also interfere with translation of the migration protein.

At the protein level, one skilled in the art can modulate the activity of the migration protein either by modifying the protein encoded by the gene as noted above or by directing the protein to be modified in vivo, for example, by providing in the protein appropriate signal or signals for cleavage or degradation by other cellular factors. Alternatively, the protein can be targeted with an activity-modulating factor such as a protein, a peptide, or an organic or inorganic co-factor. Any of these factors can, for example, occupy or obstruct an active site of the protein which is required for activity. Likewise, if the activity of the protein is natively regulated by an endogenous co-factor, an effect can be achieved by modulating the availability of the native co-factor.

One skilled in art is familiar with the techniques associated with the aforementioned alterations, including the production of any construct necessary to effect such changes. One skilled in the art also understands that changes in the primary amino acid sequence (including, e.g., substitutions, deletions, additions, inversions) may or may not alter the activity of a protein, depending upon the position and the extent of the change.

For purposes of this application a migration protein is considered active if it causes a cell that comprises the protein, or a cell that is under the influence of the protein, to migrate to any appreciable extent. A cell is "under the influence of the protein" if the cell migrates in the presence of the protein, even if the cell does not

contain the protein. *In vivo*, the cell from which the protein is secreted and its site of action remain unknown.

Non-native transgene sequences containing non-native sequences homologous to all or part of *C. elegans gon-1* can be introduced into *C. elegans* on an expressible genetic construct that contains a promoter that drives expression in a tissue that allows easy assay so that the effect or effects of those sequences on migration and other functions can be evaluated in the system. Methods for generating and selecting transgenic nematodes are well-known in the art. Transgenic animals can rescue null mutants or can suppress or enhance the activity in the reduced-activity mutants. A preferred example of a transgene sequence is a human *gon-1* homolog sequence, although any of homolog can be used. Some constructs may contain all or part of the *gon-1* coding sequences. The transgene should be appropriately expressed near the cells to be controlled by the migration protein. In *C. elegans*, the *gon-1* promoter, active in leader cells and in muscle cells, is suitable. Other promoters that can be used in *C. elegans* include the *lag-2* promoter, which drives expression in the hermaphrodite distal tip cells, and the *unc-54* promoter which drives expression in body wall muscle.

One can assay for effects of treatment with a potential modulating agent on cell migration and gonadal tube extension by comparing migration after treatment to the cell migration in either a wild-type organism or to that in an untreated, previously characterized mutant. Before treatment in the methods, if the migration protein is expressed in leader cells at wild-type levels, directed elongation of gonadal arms along a proximal-distal axis is observed. If the migration protein is expressed in muscle, on the other hand, one observes more dispersed activity, which may be important for expansion as the gonad along the dorsal-ventral and left-right axes. If a migration protein having a level of activity comparable to that of the wild type protein is expressed from a polynucleotide sequence

under control of the native *gon-1* promoter, of course, normal gonadal development is observed, as is shown in Fig. 2A. Fig. 2B shows that arm extension does not occur in *gon-1* mutants and that the gonad develops as a disorganized mass of somatic and germline tissues. Similarly, in males, the *gon-1* mutant gonad is severely disorganized and does not acquire its normal shape. Both wild-type activity and the mutant phenotype can be modified by treatment according to the methods. One can also direct the shape of a tissue or organ by introducing a transgene coding sequence under control of a promoter selected to express the transgene coding sequence in a desired tissue or cell type.

One can also assess whether a cell has the potential for migration by analyzing for example, the level of the migration protein in the cell, or the level at which the RNA encoding the migration protein is present. A diagnostic assay for the presence of active site residues in the protein can also be devised. Likewise, the presence or absence of a DNA sequence encoding an essential aspect of the protein can also be used in a diagnostic manner to assess the likelihood of cell migration.

Our finding that GON-1 is tightly regulated to achieve arm extension during gonadogenesis in C. elegans suggests that similar activities may play similar roles in the morphogenesis of organs throughout the animal kingdom. Previous *in vitro* experiments support this notion. For example, antibodies recognizing matrix metalloprotease 9 (MM9) can block branching of the ureter bud during kidney development (Lelongt et al., 1997), and inhibitors of MMPs block the invasion of endothelium cells into a fibrin matrix in assays for angiogenesis (Hiraoka et al., 1998). Based on these observations and our analysis of GON-1, we suggest that the MPT metalloproteases are critical modulators of organogenesis.

Whether the target organism contains a wild-type *C. elegans gon-1* gene, a mutant *gon-1* gene or a transgene substituted in place of *gon-1*, in whole or in part, the

system is readily used to identify other genes, proteins, drugs, chemicals or other factors that either enhance or antagonize activity.

In a method for increasing the migration of the cell, the native protein or related protein or a genetic construct encoding same can be administered to, or caused to be expressed at a high level in, the target cell. Alternatively, an enhancing factor can be provided inside or outside the target cell, as appropriate. Where it is desired to decrease migration of a targeted cell, as in the case of a tumor cell, an inhibiting factor can be added into, or the vicinity of, the targeted cell. The vicinity of the cell is defined as sufficiently close to the targeted cell so as to effect a desired change in the cell migration. If the migration protein is secreted from the cell in which it is produced, the activity of the protein can further be modulated either by preventing secretion of the protein or by interfering with the protein activity outside the cell. If the protein acts outside the target cell, the protein, an active portion thereof, or a modulating factor can be administered to the vicinity in an amount effective to modulate cell migration.

The reproductive sterility that can result from inhibited migration of developing gonadal cells under the control of an migration protein that is inactive or has reduced activity can be further exploited, for example, in a method for controlling reproduction of an organism that relies upon a migration protein during gonadogenesis. An organism for which such control would be appropriate would include *C. elegans* and other nematodes or parasites, and could include other invertebrates, as well as vertebrate species including, for example, avian, amphibian, reptilian and mammalian species.

With an appreciation for the migration proteins of the invention, normal and abnormal cell migration attributable to activity of a migration protein can be therapeutically increased or decreased. The mechanisms by which the gene

and protein are regulated can be determined by one skilled in the art and can be advantageously exploited to modulate expression of the migration protein at either the nucleic acid or protein levels.

EXAMPLES

To gain molecular insight into *gon-1* function, we cloned the gene by a combination of fine genetic mapping, mutant rescue and RNA-mediated interference. Mutations in the *gon-1* gene were finely mapped by genetic crosses with respect to markers that had already been placed on the physical map. Cosmids in the region were next tested for mutant rescue of the *gon-1* mutations. The genomic *C. elegans* sequence that includes the coding sequence of the *gon-1* gene in a plurality of exons is found on cosmids F25H8 (Accession # 69360) and T13H10 (Accession #69361); T13H10 bears most of *gon-1* and rescued the *gon-1* phenotype. The predicted open reading frames on this cosmid were tested by RNA-mediated interference to identify the transcript corresponding to *gon-1* activity. The identification of this transcript as *gon*-1 was then confirmed by subcloning and mutant rescue by a smaller region of the cosmid that contained that transcript, by RNA-mediated interference, and by identifying *gon-1* mutations in the coding region of this transcript. The positions in the migration protein that correspond to the identified mutations are indicated in Fig. 1B. We confirmed identification of F25H8.3 as *gon-1* by identifying molecular lesions for a plurality of *gon-1* alleles.

Mutants were obtained as described (Brenner, S. "The Genetics of *Caenohrabditis elegans*, Genetics 77:71-94 (1974), incorporated herein by reference. Each contained an allele of *gon-1* that maps to chromosome IV between *unc-24* and *dpy-20*, all are recessive, and all are fully penetrant for sterility. Five alleles, *e1254*, *e2547*, *q18*, *q517*, and *q518*, fail to complement the sixth allele, *e2551*, and, therefore, the mutations define a single gene. Three-factor mapping places *gon-1*(*e2551*) 0.08 map units to

the right of *elt-1* and 0.12 map units to the left of *unc-43* at position 4.44. Specifically, among Unc-43 non-Elt-1 recombinants isolated from *gon-1/ elt-1 unc-43* mothers, 8/13 carried the *gon-1* mutation.

To compare allelic strengths, we examined the penetrance of arm extension defects in homozygotes for each allele. In *gon-1(q518)* homozygotes, no arm extension was observed at 15°, 20° or 25°C. However, in homozygotes for the other *gon-1* alleles, some arms extended at least partially. By this measure, the *gon-1* alleles can be placed in an allelic series: *q518* < *e2547* ≈ *q18* < *e1254* ≈ *q517* < *e2551*. Interestingly, the weaker *gon-1* alleles have a more severe defect at lower temperature, which may reflect a cold sensitivity of GON-1 function, or of the process of arm extension itself.

The strongest loss-of-function allele is *gon-1(q518)* which is a nonsense mutation that resides in the canonical TSP1 motif; the other mutations are located in the TSP1t1-like repeats. *gon-1*(q518), the nonsense mutant located closest to the N-terminus, has the most severe effect on cell migration; nonsense mutants located closer to the C-terminus than q518 are partially defective for migration. Because the mutant phenotype for *gon-1(q518)* homozygotes is identical to that of *gon-1(q518)* hemizygotes and because *gon-1(q518)* bears a nonsense mutation predicted to remove the bulk of the GON-1 protein, this allele is likely to be a molecular null. Therefore, *gon-1(q518)* was used for analyzing the roles of *gon-1* in gonadal morphogenesis and is referred to as *gon-1(0)*.

Normally, the gonad is a tubular structure with specialized regions. By contrast, in *gon-1* mutants, the adult gonadal tissues exist as a disorganized mass with little or no tubular morphology. Specifically, neither arms nor somatic gonadal structures (e.g. uterus, spermatheca) are observed. In all cases, however, the gonads are rendered infertile by these mutations.

In *C. elegans*, mRNAs containing premature stop codons

are normally degraded by the *smg* system, but those mRNAs are stabilized in a *smg* mutant background (Anderson and Kimble, 1997). Therefore, the remaining activity of truncated GON-1 proteins should be evident in *smg-1; gon-1* double mutants. We found that *gon-1(q518)* was not suppressed in a *smg* background, whereas all four mutations in the TSP1-like repeats were suppressed. Therefore, while the GON-1(q518) mutant protein that possesses the metalloprotease domain but lacks the *bona fide* TSPt1 motif (as well as the rest of the protein C-terminally), is not capable of mutant rescue, the other truncated proteins are. The conclusion that two TSPt1-like repeats are sufficient for rescuing activity was confirmed by mutant rescue with a mini-transgene.

The lack of gonadal arms in *gon-1 (0)* mutants suggested that the leader cells, which normally govern arm extension, may be defective. To assess whether leader cells were generated during development, we first examined the gonadal cell lineages in *gon-1(0)* mutants during the first two larval stages. Normally, the somatic gonadal progenitor cells, Z1 and Z4, give rise to two leader cells, Z1.aa and Z4.pp, in hermaphrodites, and one leader cell, Z1.pa or Z4.aa, in males (Kimble and Hirsh, 1979). In hermaphrodites, these leader cells are called distal tip cells (DTC), and in males, they are called linker cells (LC). The hermaphrodite distal tip cell is both a leader cell and a regulator of germline proliferation. Kimble, J.E. and J.G. White, "On the control of germ cell development in *Caenorhabditis elegans*, Devel. Biol. 81:208-219 (1981), incorporated herein by reference in its entirety, provides guidance for a skilled artisan on the biology of distal tip cell migration. The information disclosed in that paper can be employed in determining whether an agent modulates cell migration or tissue shaping in a method of the invention.

In *gon-1(0)* hermaphrodites and males, we found that the timing and pattern of cell divisions of Z1 and Z4 and

their descendants were the same as in wild-type during L1 and L2 (data not shown). In particular, Z1.aa and Z1.pp in hermaphrodites and Z1.pa/Z4.aa in males were born at the correct time and place. To ask whether the presumptive hermaphrodite leader cells, Z1.aa and Z4.pp, had adopted the leader fate, we examined expression of a molecular marker for that fate. The *unc-5* gene encodes a netrin receptor and is essential for dorsal migration of leader cells (Leung-Hagesteijn et al, 1992). Using a reporter transgene, *unc-5::lacZ* (J. Culotti, personal communication), we found that *unc-5* expression was the same in wild-type and *gon-1(0)* animals: *unc-5* was not expressed during early larval stages, but was activated in late L3 when the DTCs normally turn dorsally during wild-type gonadogenesis.

Since the hermaphrodite leader cells, Z1.aa and Z4.pp, also control germline proliferation, we next asked if they were correctly specified for that regulatory function. To this end, we examined expression of the *lag-2* gene, which encodes the DTC signal for germline proliferation (Henderson et al., 1994). Using a reporter transgene, *lag-2::*GFP, we found that *lag-2*::GFP expression was similar in wild-type and *gon-1* gonads. Furthermore, we ablated Z1.aa and Z4.pp in *gon-1(0)* mutants and found that germline proliferation was arrested. Therefore, the hermaphrodite DTCs, Z1.aa and Z4.pp, appear to be specified correctly both as leader cells and as regulators of germline proliferation.

Since the leader cells appeared to be specified correctly in *gon-1* mutants, we next examined their ability to migrate and lead arm extension. Normally, the hermaphrodite leader cells (distal tip cells) migrate away from the center of the gonad along the anterior-posterior axis, then reflex dorsally, and migrate back. To compare leader cell migration in wild-type and *gon-1(0)* mutants, we followed their movements throughout gonadal development and at the same time measured gonadal lengths. At the

mid-L1 stage, just prior to division of the leader cell progenitors, Z1 and Z4, the length of the gonad from anterior to posterior end was 19 μm in both wild-type and *gon-1(0)* mutants. Following division of Z1 and Z4 in late L1, a small difference in gonadal length was discerned: 25 μm in wild-type vs. 22 μm in *gon-1* mutants. However, in older larvae with differentiated leader cells, the length differences were dramatic. In *gon-1(0)* hermaphrodites, the distal tip cells had moved little from their birth position and little to no gonad extension had occurred.

A similar defect is observed in males. Normally, the male leader cell (linker cell) migrates anteriorly, then reflexes and migrates to posterior end of the worm. However in *gon-1(0)* males, the linker cell failed to migrate, and little to no extension had occurred. We conclude that *gon-1* is required for leader cell migration and hence gonadal arm extension.

As we observed leader cells during gonadogenesis, we noticed that they assumed an unusual morphology. To explore this further, we examined hermaphrodite DTCs using fluorescence and thin section electron microscopy (EM). Using *lag-2*::GFP, which is expressed in hermaphrodite DTCs and reveals the extent of their cytoplasm (D. Gao and J. Kimble, unpublished), we found that the wild-type and *gon-1(0)* DTCs had dramatically different morphologies. In wild-type, the DTC was crescent-shaped with processes extending around the germ line, while in *gon-1* mutants, it was round and enlarged. Furthermore, the position of the nucleus within the DTC was variable in *gon-1* mutants, whereas in wild-type, it was located at the leading edge of the migrating cell. By EM, we confirmed the difference in morphology between wild-type and *gon-1* leader cells and also discovered a difference in subcellular organization. Whereas wild-type leader cells extend processes along the germline, *gon-1(0)* leader cells do not possess such processes. Furthermore, the plasma membrane is abnormally invaginated in *gon-1(0)* L3 leader cells, and these

membranes accumulate within the cytoplasm of older *gon-1(0)* mutants.

The lack of gonadal arms is not the only defect in *gon-1* mutants. In addition, no gonadal structures (e.g. uterus in hermaphrodites, vas deferens in males) can be discerned. One problem might have been a failure to differentiate gonadal tissues. However, we were able to identify the major somatic gonadal cell types in late L4 *gon-1(0)* mutants. To see somatic gonadal sheath cells, we used *lim-7*::GFP, which expresses Green Fluorescent Protein (GFP) in hermaphrodite sheath cells (O. Hobert, pers. comm.). In wild-type, fluorescence from *lim-7*::GFP encircled the germ cells, while in *gon-1* mutants, only irregularly-shaped patches were observed. Similarly, MH27 antibody, which stains spermathecal cells intensely (den Boer et al., 1998), was present in disorganized patches in *gon-1* mutants. Finally, cells with a typically uterine morphology were present, but no normal uterine structure was found in *gon-1* mutants. Therefore, the gonadal tissues in *gon-1(0)* mutants appear to differentiate correctly.

One simple explanation for the gross morphogenetic defects of mature *gon-1* gonads might have been that all aspects of gonadal morphogenesis are disrupted as a consequence of the defect in leader cell migration. Indeed, by killing the distal tip cells in wild-type animals, we could reproduce the *gon-1* mutant phenotype: arms did not extend and gonadal structures were grossly malformed. However, closer inspection suggests that *gon-1* has a role in gonad morphogenesis independent of leader cells.

To examine the generation of gonadal somatic structures, we removed the germ line (-GL) from *gon-1(0)* to permit formation of an essentially normal somatic gonadal primordium at the early L3 stage and we removed both leader cells (-DTCs) and germline (-GL) from wild-type hermaphrodites as a control. The control animals had no arm extension, but formed a normal somatic gonadal primordium.

A comparison of gonadal structures at the L4 stage, when they are most easily scored, revealed striking differences. While fragments of uterus were present in *gon-1*(-GL) hermaphrodites, no coherent uterus was observed. Furthermore, the *gon-1* (-GL) gonad was small, and most gonadal had extruded from the gonad proper. By contrast, an apparently normal uterus formed in the wild-type animals lacking both DTCs and germ line. Therefore, *gon-1* is required not only for arm extension, but also for morphogenesis of the uterus.

Finally, we asked whether *gon-1* functions in the development of non-gonadal tissues. We assayed embryonic viability, the overall shape of the animal, coordination of its movements, mating behavior in males, the male tail, growth rate, and entry and exit into dauer stage of the life cycle: all were normal in *gon-1(0)* mutants. The normal movement and shape of *gon-1(0)* mutants suggests that *gon-1* is not required generally for cell migration. For example, failure in migration of the CAN neuron causes the tail to wither (Forrester et al., 1998), and defects in axon migration leads to an uncoordinated (Unc) phenotype (Hedgecock et al., 1990). Furthermore, we followed the M sex myoblast and the Q neuroblasts migrations (Antebi et al, 1997)in at least five *gon-1(0)* mutants, and both were normal. We conclude that *gon-1* does not affect cell migrations generally and, furthermore, that *gon-1* does not affect the development of non-gonadal cells, tissues or organs. Finally , we examined the non-gonadal tissues in *gon-1* mutants that had been operated during L1 to remove Z1-Z4, the four gonadal progenitor cells. This experiment was done, because the disorganized gonadal tissues in *gon-1(0)* hermaphrodites often cause the animal to explode during adulthood, preventing examination of their non-gonadal tissues at this stage. Although these gonadless *gon-1* adults had no gross defects, we observed a reproducible vacuolization in the body wall with differential interference contrast microscopy, which was

not seen in similarly treated wild-type animals. However, it must be emphasized that this defect has no apparent developmental consequences. Given the dramatic effects of *gon-1* on gonadogenesis, we suggest that the major role of *gon-1* in development is to control the shape of the gonad.

The wild-type *C. elegans gon-1* sequence is shown in SEQ. ID. NO. 1. The protein encoded by SEQ. ID. NO. 1 is shown in full in SEQ. ID. NO. 2 and in part in comparative Fig. 1C.

PROPHETIC EXAMPLE

A target organism that contains a migration protein is treated with one or more potential modulators of migration of a developing gonadal cell. The organism is preferably a nematode, and is more preferably *C. elegans*. The potential modulating agent is administered in an amount typical of any additive to a culture, preferably at a level of several nanograms to several micrograms per milliliter. The organism can contain a native migration protein or a variant form of a native migration protein, or can express a migration protein from a transgene that can be delivered to the organism in a manner known to those skilled in the art. The protein can also be a chimeric protein expressed from a transgenic polynucleotide that comprises sequences from at least one of the foregoing polynucleotides.

Upon examination, it is observed that one can rescue migration in a target that lacks the migration protein by administering an exogenous polynucleotide that encodes a migration protein. In a target that contains a migration protein, one can also identify administered agents that increase or decrease the migration of a developing gonadal cell. One can also treat the genetic material of the target organism using standard methods and treatments and can then identify genetic changes that increase or decrease migration of developing gonadal cells.